CHAPTER 13
SCALING

The MEA tests were designed to measure student performance against the learning standards described in the curriculum frameworks. Consistent with this purpose, primary results on the MEA tests are reported in terms of performance levels that describe student performance in relation to these established state standards. There are four performance levels: Exceeds the Standards, Meets the Standards, Partially Meets the Standards, and Does Not Meet the Standards, as described in Chapter 12. Students received a separate performance level classification (based on scaled score) for each test. School and district level results were reported as the number and percentage of students who attained each performance level at each grade level tested.

In addition to performance levels, MEA results are reported as scaled scores. Scaled scores in each content area range from 501 to 580. Scaled scores supplement the MEA performance level results by providing information about the position of a student's results within a performance level. School- and district-level scaled scores are calculated by computing the average of student-level scaled scores.

TRANSLATING RAW SCORES TO SCALED SCORES (SCALING)

For Reading, Writing, Mathematics, Science, and Social Studies students' raw scores, or total number of points, on the MEA tests are translated to scaled scores using a process called scaling**.** Scaling simply converts raw points from one scale to another. Converting from raw scores to scaled scores does not change the rank ordering of students, give more weight to particular questions, or change students' performance level classifications.

Linear scaling parameters were determined so the minimum scaled score for Partially Meets the Standards was 521, the minimum scaled score for Meets the Standards was 541, and the minimum scaled score for Exceeds the Standards was 561. This was done by solving two linear equations relating the raw threshold scores to these predetermined scaled score values. The resulting functions that translate raw scores to scaled scores are:

$S = m_1 r + b_1$	if $r > P$, and
$S = m_2 r + b_2$	if $r < P$

where S is the scaled score, r is the raw score, and P is the Meets the Standards threshold. The values of the ms and the bs are shown in Table 13-1.

For Health and Visual and Performing Arts, scaled scores were not computed from raw scores but from estimates of students' latent traits (θ) based on models of Item Response Theory (IRT). An IRT calibration was performed for each of Health and Visual and Performing Arts tests for each of the three grade levels using the two-parameter logistic models for multiple-choice items and the graded response model for open response items. Each of these models expresses examinees, tendencies to achieve certain scores on items contributing to a scale as a function of a parameter that is not directly observed (i.e., latent). An estimate of the latent trait ($\hat{\theta}$) was obtained for each student for each test. A student's $\hat{\theta}$ is based only on the items that he/she took.

Linear scaling parameters for Health and Visual and Performing Arts were obtained using the same procedure used for Reading, Writing, Mathematics, Science, and Social Studies described above. The resulting functions that translate $\hat{\theta}$s to scaled scores are:

where *S* is the scaled score, $\hat{\theta}$ is the latent trait estimate, and *P* is the Meets the Standards threshold. The values of the *m*s and the *b*s for Health and Visual and Performing Arts are also shown in Table 13-1.

Table 13-1
Transformation Constants Used to Compute Scaled Scores

Grade	Subject Area	Transformation Constants			
		m_1	b_1	m_2	b_2
4	Reading	1.55	488.66	1.61	486.70
	Writing	2.47	495.08	2.31	498.11
	Mathematics	1.97	489.51	1.93	490.55
	Science	3.15	454.99	1.47	500.76
	Social Studies	2.91	467.12	2.54	476.66
	Health	19.68	533.95	10.13	537.37
	Visual and Performing Arts	8.21	534.14	11.40	531.48
8	Reading	1.69	484.95	1.67	485.63
	Writing	2.19	501.32	2.79	490.60
	Mathematics	1.55	503.17	1.64	500.92
	Science	2.59	473.59	2.01	488.83
	Social Studies	2.49	482.78	2.16	489.90
	Health	12.29	537.45	10.74	537.89
	Visual and Performing Arts	9.39	534.99	14.29	531.86
11	Reading	1.844	472.59	1.45	486.90
	Writing	2.92	482.21	2.49	490.85
	Mathematics	1.71	499.13	1.73	498.77
	Science	2.48	475.91	1.57	499.83
	Social Studies	2.07	497.53	2.43	490.01
	Health	13.89	536.26	10.78	537.32
	Visual and Performing Arts	5.12	536.29	14.81	527.37

After the transformation constants were applied, scores were rounded to the nearest even integer. Transformed scores below 502 were reported as 502; transformed scores above 580 were reported as 580. The slopes for Health and Visual and Performing Arts are substantially higher than the slopes for other content areas because these tests are significantly shorter than the others.

In any given year, test form difficulty and rounding might lead to some scaled scores between 502 and 580 not being attainable. For the 1998 MEA, for all subjects and grades, 502 was an obtainable value. Table 13-2 reports the highest and lowest attainable scaled scores on the 1998 MEA.

Table 13-2
Minimum and Maximum Obtainable Scores on the 1998 MEA

Grade	Subject Area	Raw Score		Scaled Score	
		Minimum	Maximum	Minimum	Maximum
4	Reading	0	53	502	570
	Writing	0	30	502	570
	Mathematics	0	41	502	570
	Science	0	41	502	580
	Social Studies	0	39	502	580
	Health	0	28	502	580
	Visual and Performing Arts	0	28	502	580
8	Reading	0	52	502	574
	Writing	0	30	502	568
	Mathematics	0	41	502	566
	Science	0	41	502	580
	Social Studies	0	41	502	580
	Health	0	28	502	580
	Visual and Performing Arts	0	28	502	580
11	Reading	0	53	502	570
	Writing	0	30	502	570
	Mathematics	0	41	502	570
	Science	0	41	502	578
	Social Studies	0	39	502	578
	Health	0	28	502	580
	Visual and Performing Arts	0	28	502	580